The Earth System Model

Mark Schoeberl, Peter Hildebrand, Richard Rood
NASA, Goddard Space Flight Center
Robert Ferraro, Carol Raymond
NASA, Jet Propulsion Laboratory

The Earth System Model

"Prediction is the true test of our knowledge" - John Dutton

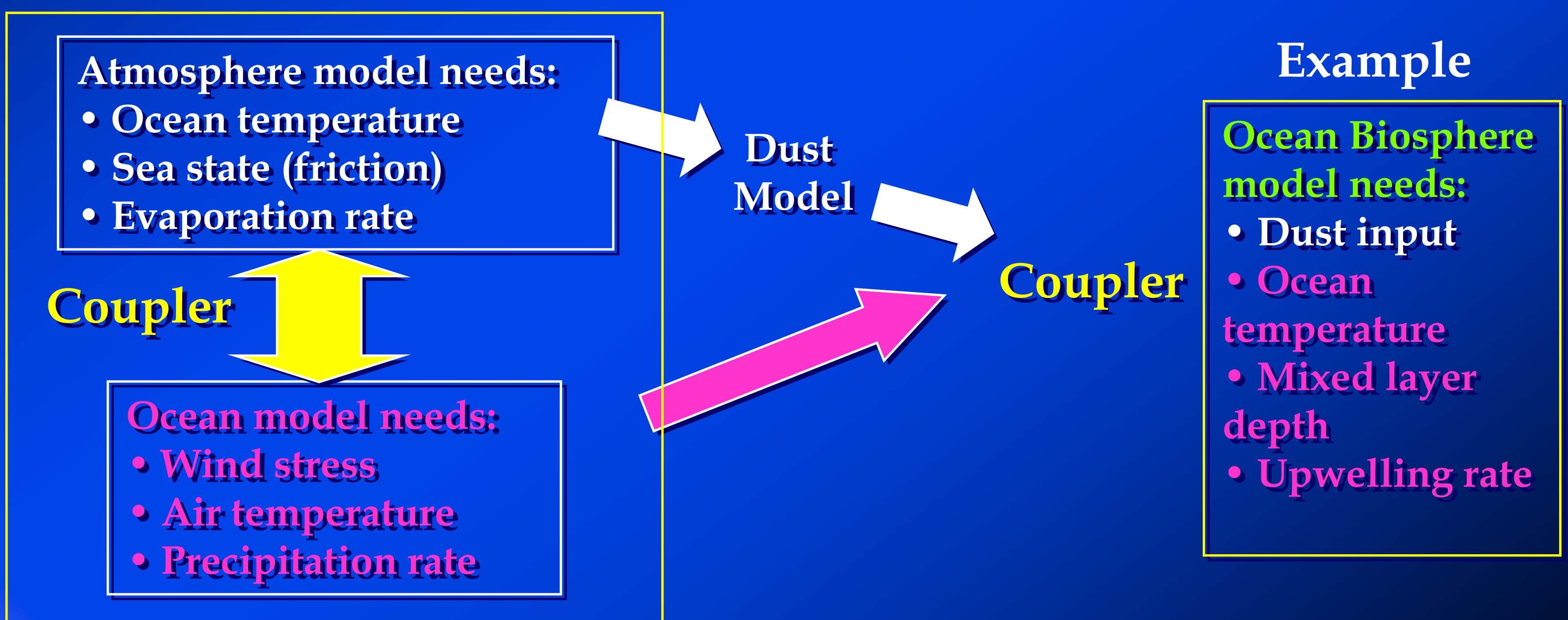
An Earth System Model is the coupling of separate model components in such a way as to describe the interactions between different processes.

Why build an Earth System Model?

- *To provide useful and accurate predictions of processes that are too complex for single system models*
  - *e.g. effects of air-sea interaction on climate and weather forecasts*
- *To provide an assessment of the importance of feedbacks between different processes*
  - *e.g. seasonal land cover change on climate*
- *To extend prediction capabilities into new regimes*
  - *e.g. biosphere impacts of climate change*

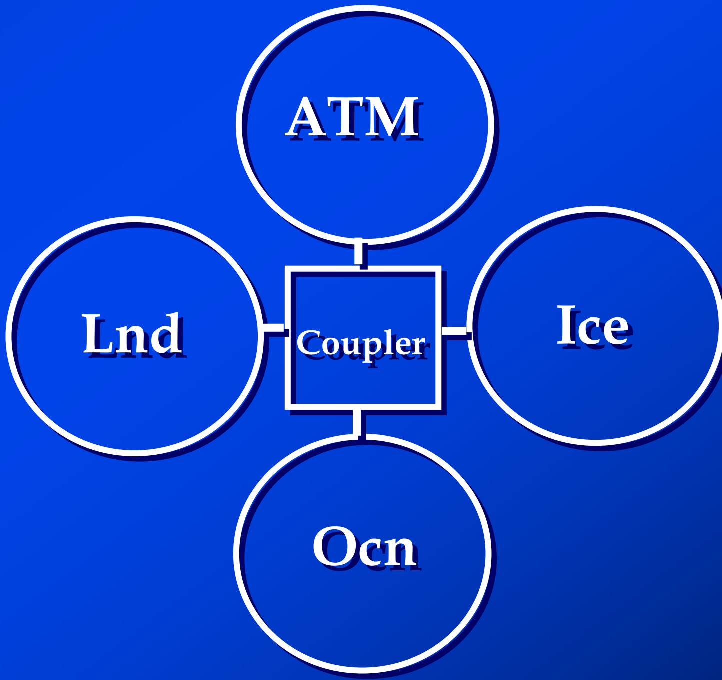
ES Model Evolution

- ES models begin by coupling two or more models together.
  - The starting point is often the atmospheric model.
    - Ex/ Atmospheric-ocean model coupling
    - Add biosphere model coupling



ES Model Framework

- Comprehensiveness
  - Number of sub-models
- Coupling
  - The coupler (or supervisor) is the heart of the system and coordinates the models
  - Each model retrieves what it needs from the other components, while placing output there for other models.
- Modularity
  - The ESMF or PRISM standardizes model interfaces within and between the models
  - Community participation in development is facilitated by the modules within an ESM framework
- Scalability
  - As model resolution increases parameterization schemes will need to scale
    - e.g. bulk convection to cloud resolving systems



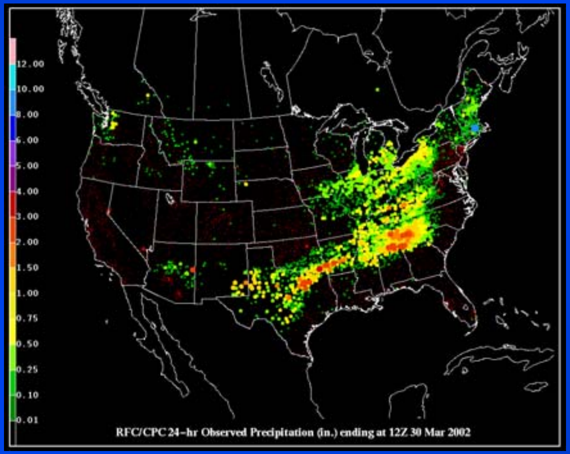
Example: the NCAR CCSM Framework

ESM Computing Issues

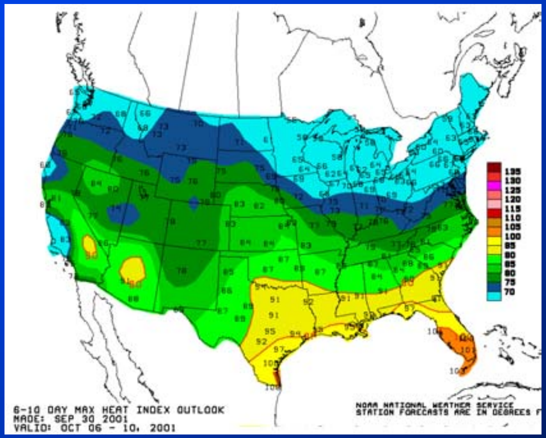
- Spatial Resolution
  - The higher the spatial resolution the more improved the results
  - Higher resolution models need shorter time steps (Courant limit) or need to use semi-implicit solvers (… i.e., more computer time)
  - Linkage between horizontal resolution and vertical resolution (~1/100, f/N in atmospheric models)
- Physics Packages
  - Become increasingly complex as science improves (e.g. cloud models)
- Diagnostics
  - Storage and visualization stress IO bandwidth and storage capacity
- Assimilation
  - Ingest data organization
  - Data preparation
    - Need to compare model output with observations for quality check
  - Optimized initial state estimation
- Predictions
  - Ensemble runs

Weather Prediction

Precipitation



Heat Index



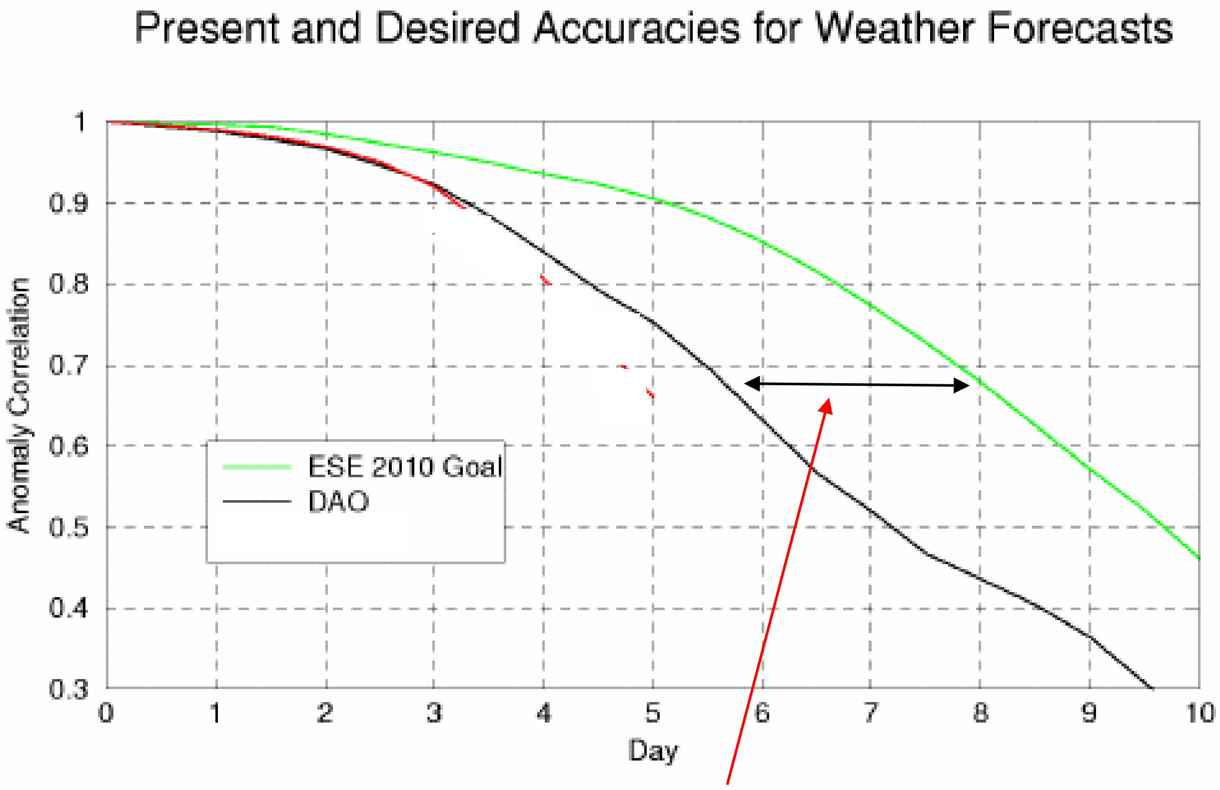
ESE Prediction Goals for Weather Prediction

Today's Capability

- 3-day forecast at 93%
- 7 day forecast at 62%
- 3 day rainfall not achievable
- Hurricane landfall+/- 400 Km at 2-3 days
- Air Quality day by day

2010+ Capability

- 5 day forecast at >90%
- 7-10 day forecast at 75%
- 3 day rainfall forecast routine
- Hurricane landfall+/- 100 Km at 2-3 days
- Air Quality forecast at 2 days

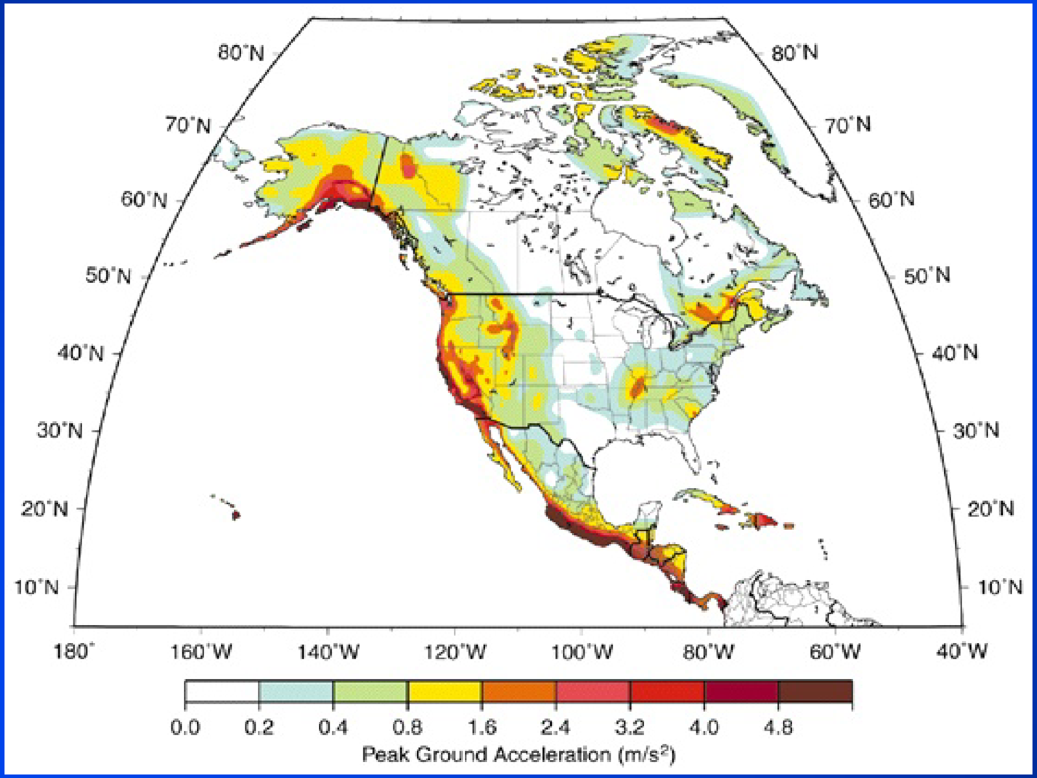


2.5 day skill improvement
(How are we going to do this?)

Requirements for Weather

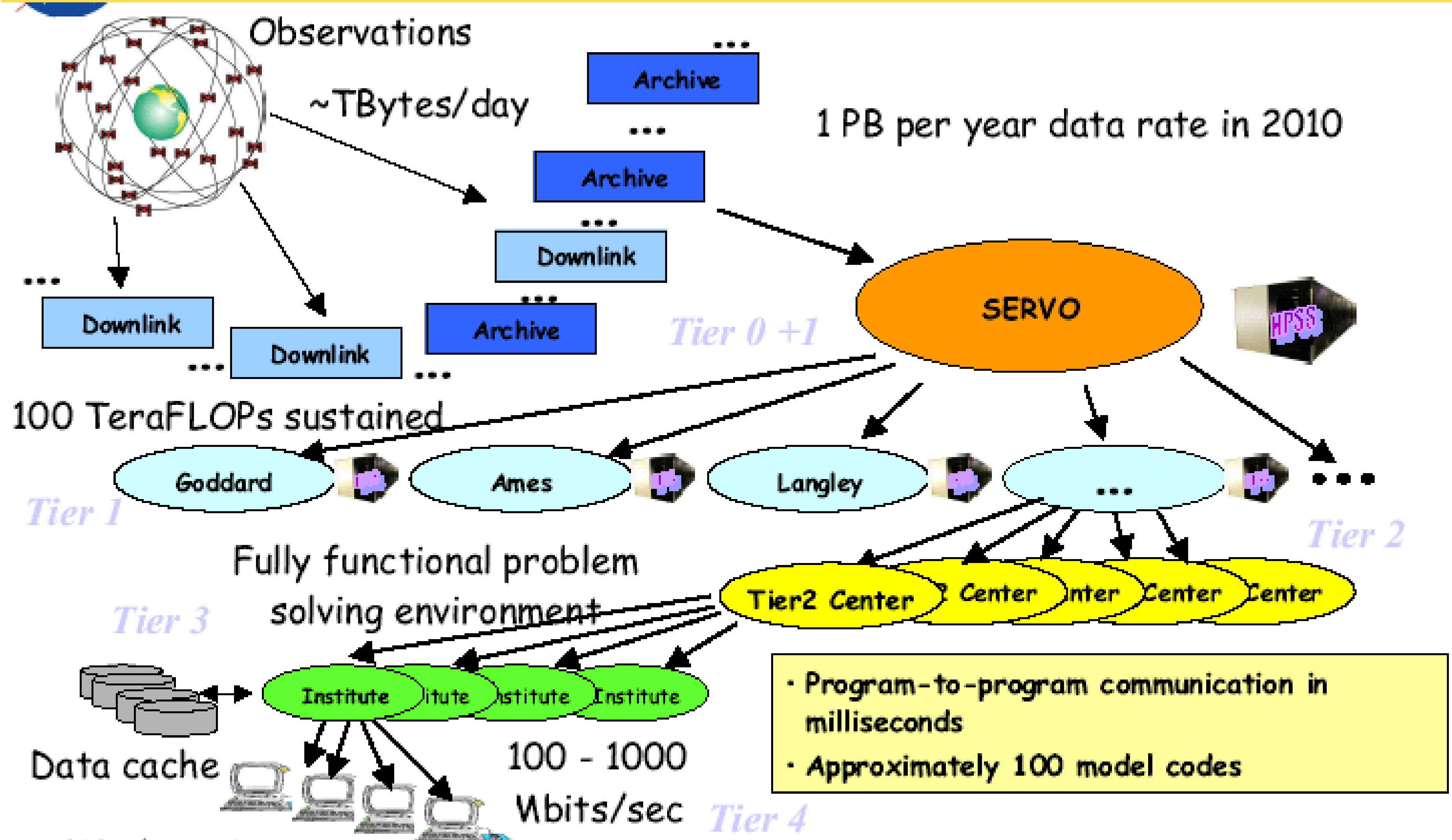
	2002 System	2010+ System	
Resolution • Horizontal • Vertical levels • Time step • Observations ○ Ingested ○ Assimilated	100 km 55 30 minutes 10^7 / day 10^5 / day	10 km 100 6 minutes 10^{11} / day 10^{8}/ day	
System Components:	Atmosphere Land-surface Data assimilation	Atmosphere, Land-surface, Ocean, Sea-ice, Next-generation data assimilation Chemical constituents (100)	
Computing: • Capability (single image system) • Capacity (includes test, validation, reanalyzes, development)	10 GFlops 100 GFlops	Must Have 20 TFlops (2000x) 400 TFlop (4000x)	Important 50 TFlops 1 PFlops
Data Volume: • Input (observations) • Output (gridded)	400 MB / day 2 TB / day	1 PB / day 10 PB / day	
Networking/Storage • Data movement ○ Internal ○ External • Archival	 4 TB / day 5 GB / day 1 TB / day	 20 PB / day 10 TB / day 10 PB / day	

Solid Earth Modeling



Seismic Hazard Map

Solid Earth Research Virtual Observatory (SERVO)



- Plug and play composing of parallel programs from algorithmic modules
- On-demand downloads of 100 GB in 5 minutes
- 10^6 volume elements rendering in real-time

Computing and Data Storage Requirements

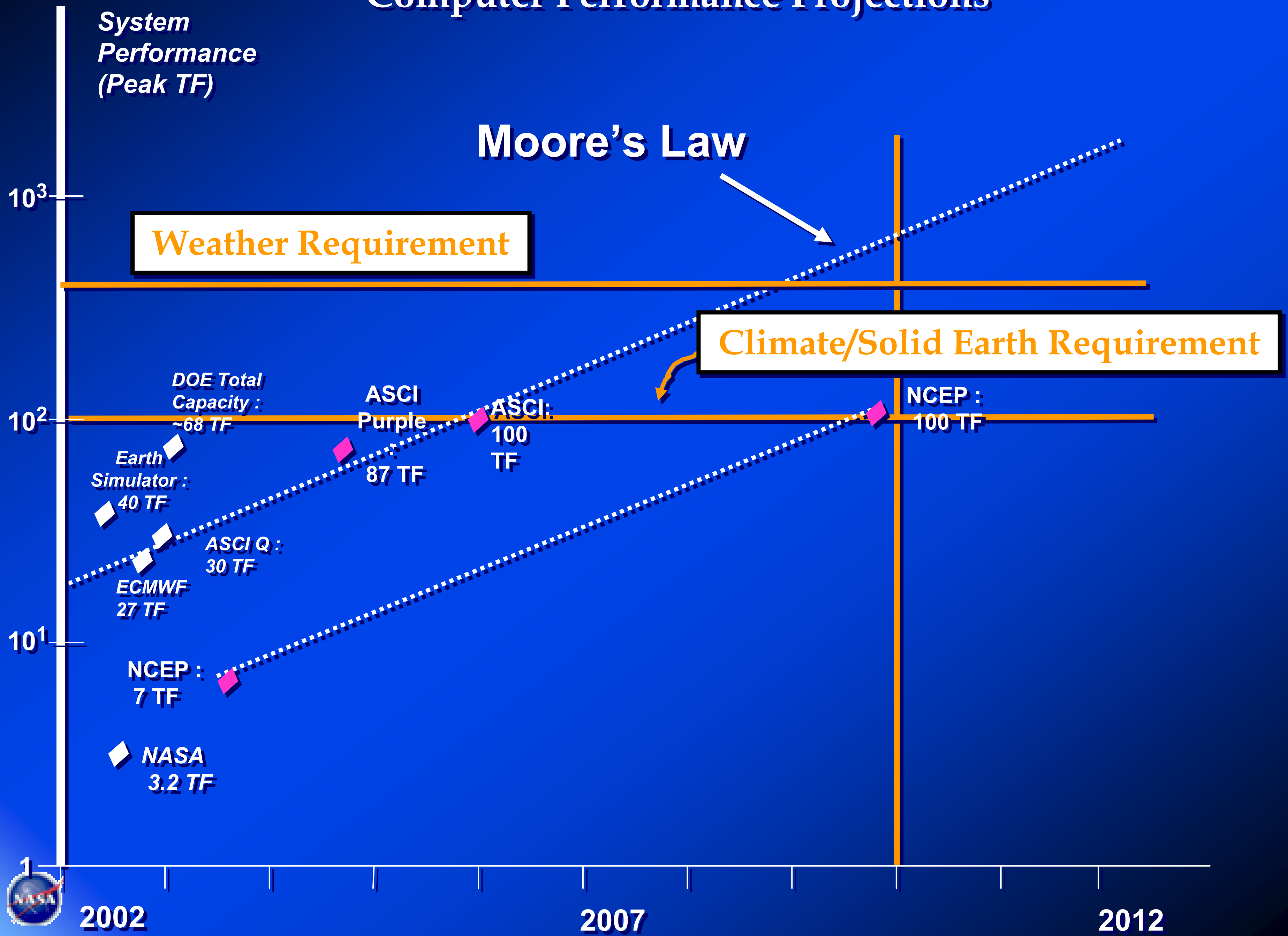
A. Computing Platform Throughput Required

	Stressing Model	Single Image Throughput	Estimated Capacity Required
Weather	10 Day Forecast Atmosphere:10 km horizontal, 100 levels vertical 10^{11} observations	20 Tflops	400 Tflops
Climate	S-I Prediction Atmosphere: 25 km horizontal Ocean: 6 km horizontal	5 Tflops	100s Tflops
Solid Earth	Earthquake Fault Slip 16M finite elements 100k boundary elements	2 Tflops	10s – 100 Tflops
Sustained Throughput and Capacity Requirements			

– Single application requirements derived from current performance extrapolated by required resolution increase

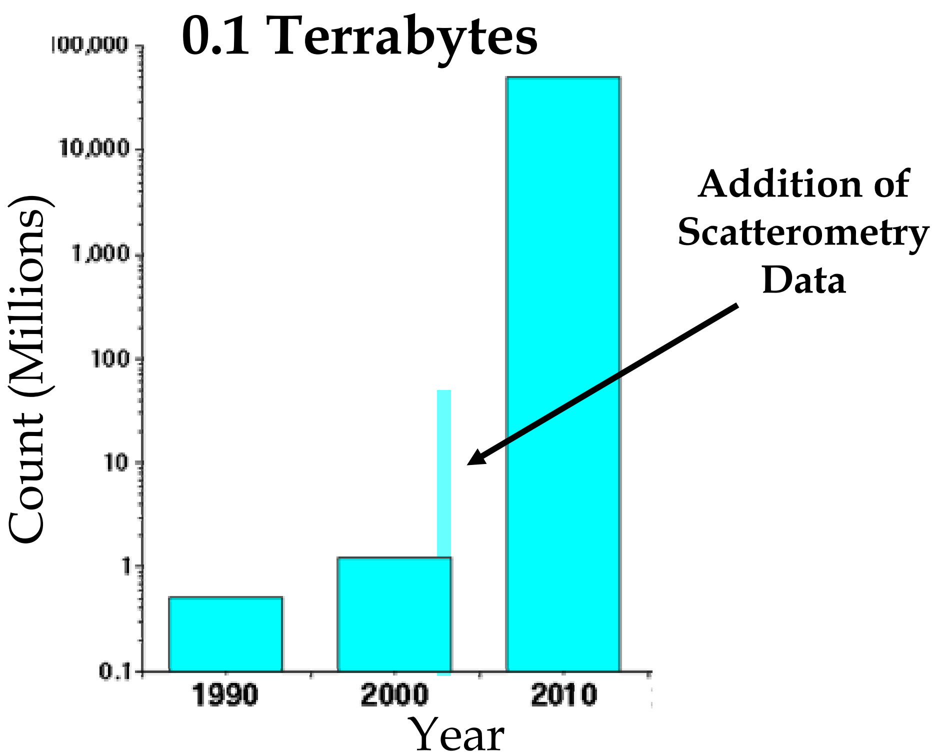
– Capacity requirements are based on current experience scaled up to the 2010 strawman environments

Computer Performance Projections

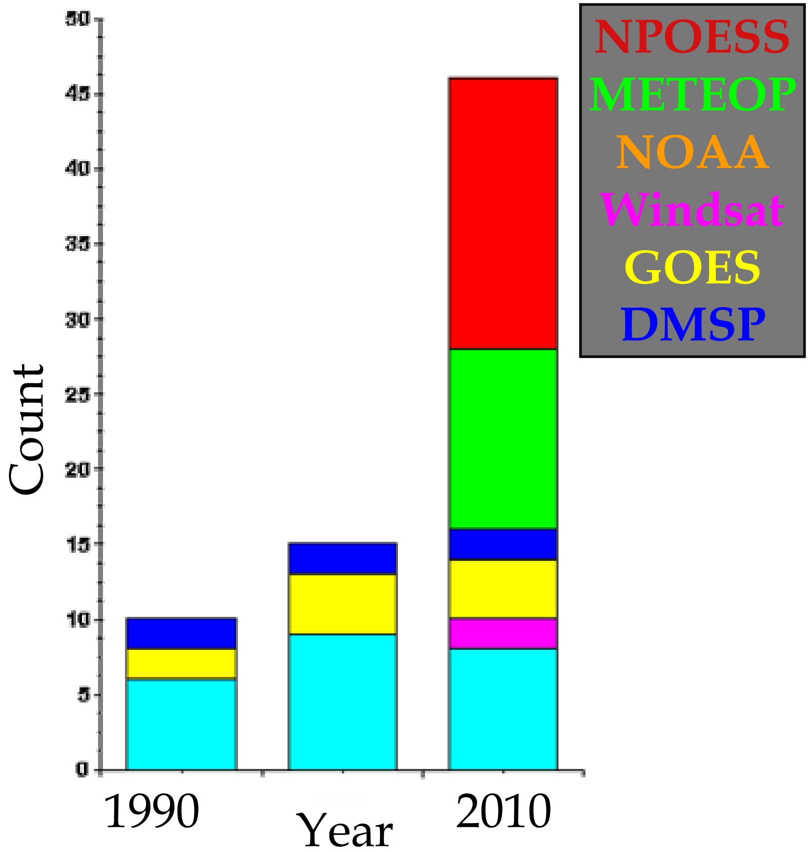


5-Order Magnitude Increase in Satellite Data Over 10 Years from NOAA Weather Platforms

Daily Upper Air Observation Count



Satellite Instruments by Platform



Data Management Requirements

	Observational Data	Access Modes Rates	Output Data	Storage Term/Re-access Mode
Weather Forecast	1 TB/day Multiple Sources Continuous	Streamed input 20 GB/s	10 PB/day – Archival 10 TB/day – external distribution	Medium – Long Catalogued
Climate Modeling	10s of GB from archival sources	Data archive request 2 GB/s (latency tolerant)	100s TB/day	50% Short term - Immediate analysis 50% Medium term - Catalogued
Solid Earth Research	100s of GB/day Distributed sources	Distributed archives – low latency access	1 PB/day – ingested into distributed archives	Medium – Long Catalogued access

– Data volume is expected to be overwhelming and heterogeneous in format

Model output data management is the problem

– Current practice does not scale to these volumes

– Data storage expected to be geographically distant from data consumers

– Uniform, seamless identification, indexing, and access methods required

Data Volume Issues

In the far future a completely different paradigm will be required in which data volumes are vastly reduced:

- The vast amounts of data collected by sensors can be pre-processed and lossey compressed (1/100)
- Data assimilation tasks might be off-loaded toward the observational nodes of the system

- Model output is also not shipped around, but is summarized into manageable, user-specific chunks that are passed off to users in image format.

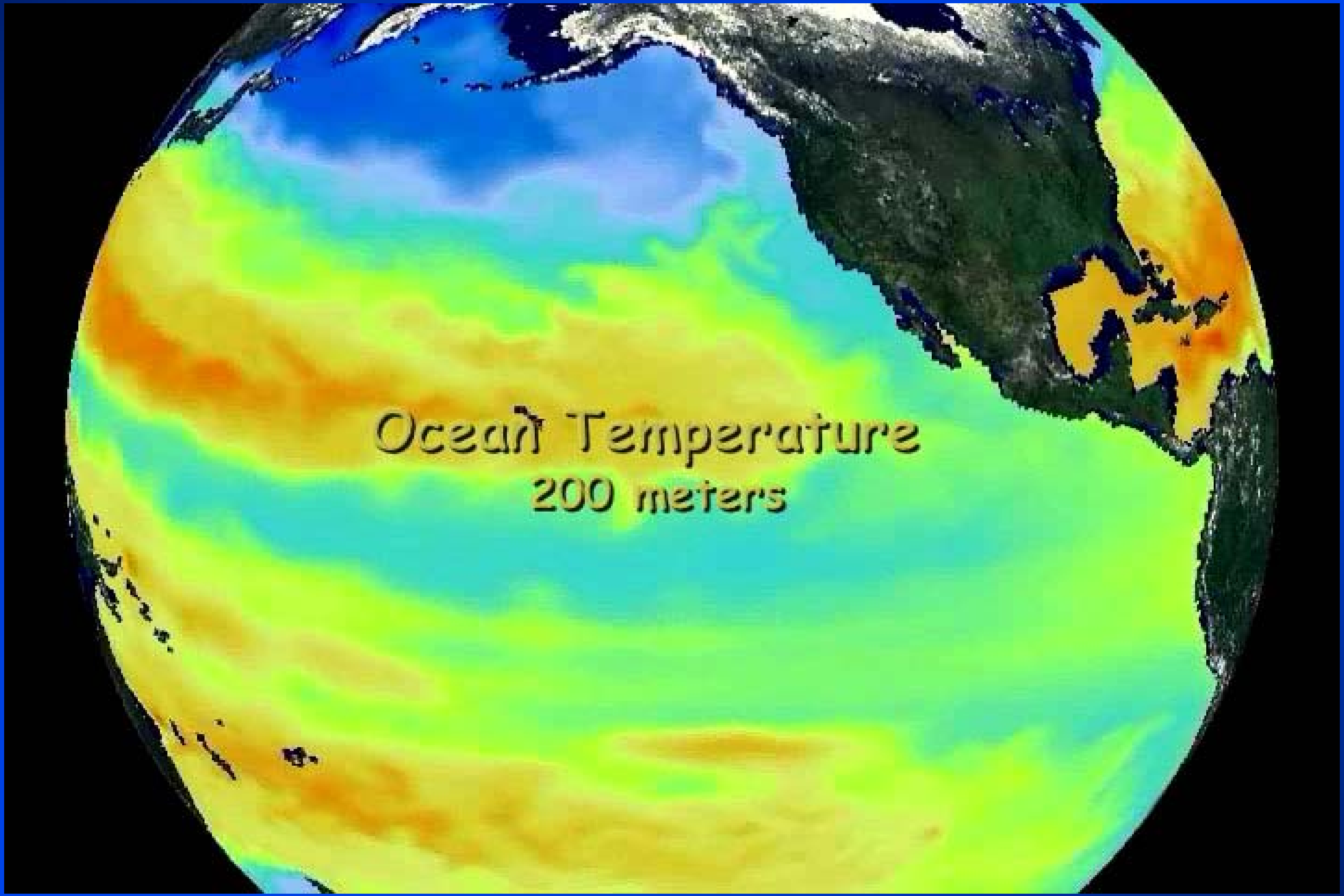
- New types of data organization methods will be needed: Users will need to be able to “google” the data they need

The Potential of the ESM
(two examples)

NSIPP* Model

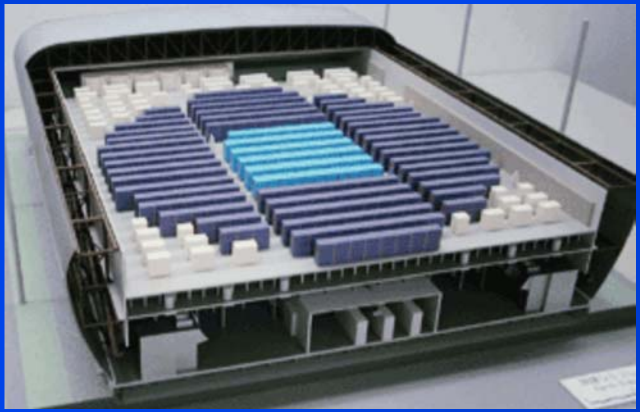
- **The goal of NSIPP is to develop an assimilation and forecast system to improve the prediction of ENSO and other major seasonal-to-interannual signals.**
  - NSIPP couples ocean, land and atmospheric models
  - NSIPP models have moderate resolution
    - Atmosphere - 2 x 2.5^o up to 10 mb
    - Ocean - 1/3 x 5/8^o - 27 layers
    - Land - tiles at 2 x 2.5^o
  - NSIPP is capable of data assimilation as well as free running climate predictions
  - NSIPP operates on the Compaq Parallel computer (1392 processors, 3.2 Tf)
  - Runtime: 64 processors, - 12 hours/year

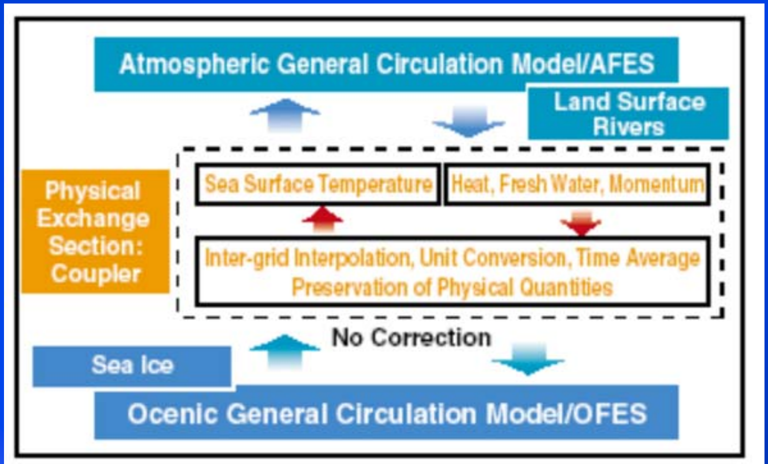
* NASA Seasonal-to-Interannual Prediction Project


Ocean Temperature
200 meters

Japanese Earth Simulator

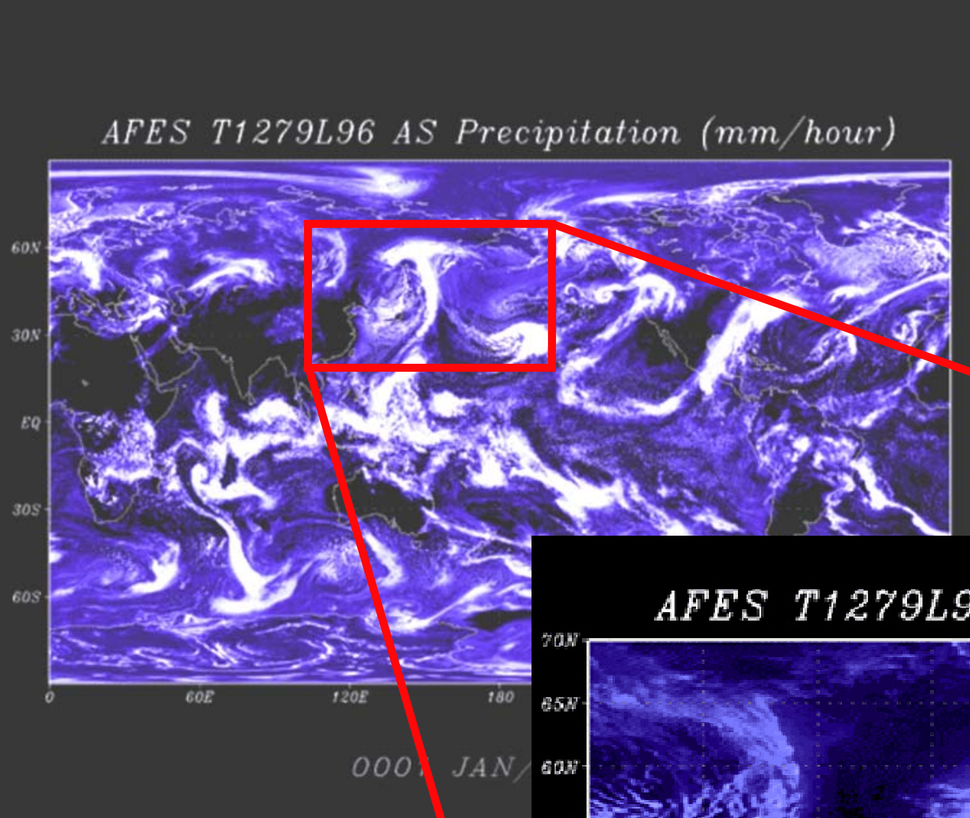
- The ES is a highly parallel vector supercomputer system of the distributed-memory type, and consisted of 640 processor nodes (PNs) the theoretical performance of ES is 40Tflops.
- There are several models being developed separately
  - AFES (atmospheric) T1279L96
  - OFES (high resolution ocean)
- ES Models are being coupled!



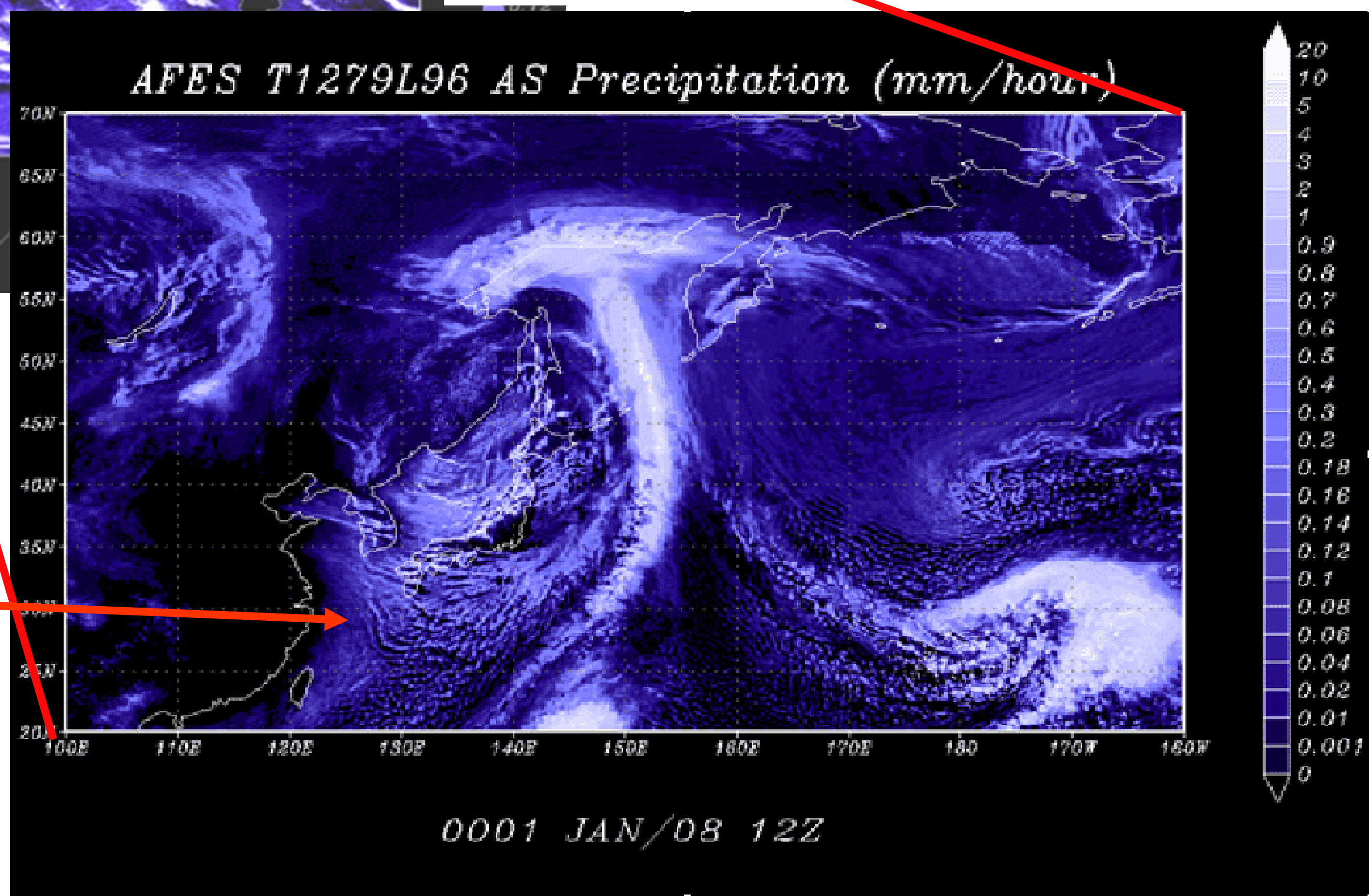


OFES

QuickTime™ and a DV/DVCPRO - NTSC decompressor are needed to see this picture.



Japanese Earth Simulator
T1279L96 spectral model
(10x10 km with 96 levels)



Note realistic cloud banding

Conclusions

- Earth System Models are the natural extension of medium range climate models
- Current ESM's are designed around the coupling of separate atmospheric, ocean, land and cryosphere models.
- There is a computer resource tension between higher spatial resolution and the desire for more coupling and more sub-models.
- Major investments in computing resources will be required to reach the 2010 capability requirements
- Focused attention on constructing and validating Earth System Models will be required to vet the science in coupled models and turn the ES into a predictive system.
- Preliminary results from various efforts is promising.

END

NASA ESE Medium-to-Long-Term Climate Goals for 2010: 10-year experimental prediction

Single Image	aggregate throughput
1.3TF	65TF
5GF	13GF

Req'd turnaround: 1000 days/day

Required Scientific Capabilities

- Coupled atmosphere-land-ocean-ice-airchem-carbon models
- Resolve interannual-to-decadal variability
- use the same model, coarser resolution for longer global change integrations (~100 years)
- many small ensembles to explore parameter space
- include stratosphere, land cover/use changes

HARDWARE

Improve scaling: shared memory utilization, I/O, internode communication

single processor optimization

Network bandwidth

2002

Lack of computer resources for required turnaround

2°X53L atm 4°X13L ocn
10 member ensembles

2005

SOFTWARE TOOLS

- Resources for code optimization

DATA MANAGEMENT

- Manually generated catalogs & metadata

2010

SOFTWARE TOOLS

- Transition to ESMF
- Analysis and visualization
- Data management software conforming to community code & data structure standards

DATA MANAGEMENT

- Fast access to remote 30TB

1°X100L atm, 1/2°X50L ocn; 40 tracers (10X computer time; 5-member ensembles
Generate 50TB/day

Decade-century integrations

Experimental decadal predictions; Global change assesments

Conceptual Evolution of ES Models